

AI chatbots not yet ready for clinical use

Year: 2023 | Citations: 129 | Authors: Joshua Au Yeung, Zeljko Kraljevic, Akish Luintel, Alfred Balston, Esther Idowu

Abstract

As large language models (LLMs) expand and become more advanced, so do the natural language processing capabilities of conversational AI, or “chatbots”. OpenAI’s recent release, ChatGPT, uses a transformer-based model to enable human-like text generation and question-answering on general domain knowledge, while a healthcare-specific Large Language Model (LLM) such as GatorTron has focused on the real-world healthcare domain knowledge. As LLMs advance to achieve near human-level performances on medical question and answering benchmarks, it is probable that Conversational AI will soon be developed for use in healthcare. In this article we discuss the potential and compare the performance of two different approaches to generative pretrained transformers—ChatGPT, the most widely used general conversational LLM, and Foresight, a GPT (generative pretrained transformer) based model focused on modelling patients and disorders. The comparison is conducted on the task of forecasting relevant diagnoses based on clinical vignettes. We also discuss important considerations and limitations of transformer-based chatbots for clinical use.